CompCytogen 10(4): 483–504 (2016)
doi: 10.3897/CompCytogen.v10i4.7582
http://compcytogen.pensoft.net

RESEARCH ARTICLE

COMPARATIVE Cytogenetics
A peer-reviewed open-access journal
International Journal of Plant & Animal Cytogenetics, Karyosystematics, and Molecular Systematics

Comparative analysis of the circadian rhythm genes *period* and *timeless* in *Culex pipiens* Linnaeus, 1758 (Diptera, Culicidae)

Elena V. Shaikevich[1], Ludmila S. Karan[2], Marina V. Fyodorova[2]

1 *Vavilov Institute of General Genetics, Gubkin str., 3, 119991, Moscow, Russia* **2** *Central Research Institute of Epidemiology, Novogireevskaya 3a, Moscow, 111123 Russia*

Corresponding author: *Elena V. Shaikevich* (elenashaikevich@mail.ru)

Academic editor: *V. Lukhtanov* | Received 22 December 2015 | Accepted 24 August 2016 | Published 10 October 2016

http://zoobank.org/29396044-9A3D-4716-8FF4-D4F7F3CDA13D

Citation: Shaikevich EV, Karan LS, Fyodorova MV (2016) Comparative analysis of the circadian rhythm genes *period* and *timeless* in *Culex pipiens* Linnaeus, 1758 (Diptera, Culicidae). Comparative Cytogenetics 10(4): 483–504. doi: 10.3897/CompCytogen.v10i4.7582

Abstract

Nucleotide sequences of the circadian rhythm genes, *period* and *timeless*, were studied for the first time in mosquitoes *Culex pipiens* Linnaeus, 1758. In this work we evaluated variations of the studied genome fragments for the two forms of *C. pipiens* (forma "pipiens" – mosquitoes common for aboveground habitats, forma "molestus" – underground mosquitoes). We compared *C. pipiens* from Russia with transatlantic *C. pipiens* and subtropical C*ulex quinquefasciatus* Say, 1823. Our results show that intraspecies variability is higher for the gene *period* than for the gene *timeless*. The revealed substitutions in nucleotide sequences and especially in amino acid sequences grouped the individuals of the two forms into distinct clusters with high significance. The detected fixed amino acid substitutions may appear essential for functioning of the circadian rhythm proteins in *C. pipiens*, and may be correlated with adaptations of the taxa within the group *C. pipiens*. Our results suggest that natural selection favors fixed mutations and the decrease in diversity of the genes *period* and *timeless* in mosquitoes of the *C. pipiens* f. "molestus" compared with the *C. pipiens* f. "pipiens", is probably correlated with adaptive features of *C. pipiens* f. "molestus". The studied genome regions may be considered as promising molecular-genetic markers for identification, population and phylogenetic analysis of similar species and forms of the *Culex pipiens* complex.

Keywords

Culex pipiens, circadian rhythm genes, *period*, *timeless*, natural selection

Copyright *Elena V. Shaikevich et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

The *Culex pipiens* Linnaeus, 1758 complex considered by some authors as a 'polytypic species' includes up to seven morphologically identical or very similar forms (Harbach et al. 1984, 1985, Vinogradova 2000). By the second half of the 20th century, the taxonomic status of these forms changed several times from species to subspecies and back. At present only two species, *Culex pipiens* Linnaeus, 1758, and *Culex quinquefasciatus* Say, 1823 have been left within the *Culex pipiens* complex based on morphological similarity (Harbach 2012). Both species are known as bridge-vectors of West Nile and Saint Louis encephalitis flaviviruses, the etiological agents of dangerous human diseases (Vinogradova 2000). The medical significance of the *Culex pipiens* complex generates much interest in its studies, including taxonomy.

Only one species of the *Culex pipiens* complex, *C. pipiens*, has been found in Russia. This species includes two forms, *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus", originally described as distinct species (Harbach et al. 1984, 1985). *C. pipiens* forms designation is provided in accordance with the rules of International Code of Zoological Nomenclature (http://www.iczn.org/iczn/index.jsp). The two forms are morphologically identical, but have notably distinct biological features. The mosquitoes *C. pipiens* f. "pipiens" are anautogenous (females require a blood meal to mature each egg raft), mate in swarms, oviposit in a wide variety of natural and manmade habitats, feed preferentially on avian hosts and enter diapause to overwinter (Vinogradova 2000). In contrast *C. pipiens* f. "molestus" are autogenous (females oviposit the first egg raft without bloodmeal), develop without winter diapause in urban flooded basements and tunnels, feed preferentially on mammal hosts and are able to mate in a confined space. The specific features of reproduction and development of the two forms has resulted in their spatial isolation in moderate climate areas, suggesting genetic isolation. This suggestion is confirmed by the isoenzyme analysis of autogenous and anautogenous populations of *C. pipiens* from England (Byrne and Nichols 1999), Russia (Lopatin 2000) and Germany (Weitzel et al. 2009) as well as by study of populations from Europe with CQ11 assay (Bahnck and Fonseca 2006). The results of these investigations showed that in these regions the forms are genetically distinct, with no or poor gene flow between populations of different forms. However, in the Mediterranean area, in N Africa and the Middle East, both autogenous and anautogenous specimens develop in the same pools. These populations display highly variable autogeny rates, from 10-90% in Egypt (Gad et al. 1995) to 4–55% in Israel (Nudelman et al. 1988), and both autogenous and anautogenous females were encountered in the progenies of autogenous or anautogenous female parents (Gad et al. 1995). Consequently, the question of divergence of the two forms in moderate climates remains still unclear.

Among the specific behavioral/physiological traits which remained up to now the important criteria for defining populations of *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus", differences in mating behavior are under the special interest. Mating ac-

tivity of *C. pipiens* f. "pipiens" is restricted within the crepuscular period when males aggregate in swarms where they copulate with virgin females attracted to a swarm (Ivanov 1984, Fyodorova and Serbenyuk 1999, Vinogradova 2000). In contrast, males of *C. pipiens* f. "molestus" never swarm and have irregular locomotor and mating activity (Shinkawa et al. 1994). Such temporal differences in mating activity may represent the temporal isolation between two forms.

In insects the rhythms of mating activity are controlled by endogenous circadian clocks, which are under genetic control (Konopka and Benzer 1971, Sakai and Ishida 2001, Tauber et al. 2003). The differences in the daily timing of mating activity are documented in many sympatric sibling insect species, e.g in tephritid fruit flies (An et al. 2002, 2004), in *Drosophila* Fallen, 1923 species (Sakai and Ishida 2001, Tauber et al. 2003), sand fly species (Rivas et al. 2008), in *Nasonia* Ashmead, 1904 wasps (Bertossa et al. 2013), in cricket species (Fergus and Shaw 2013). Intra-specific differences in the rhythms of mating activity were revealed also between populations or strains, e.g. in fly *Bactrocera cucurbitae* Coquilletl, 1849 (Fuchikawa et al. 2010) and mosquitoes of *Anopheles cruzii* Dyar and Knab, 1908 complex (Rona et al. 2010).

Clock genes, especially *period* and *timeless*, play an essential role in regulation of mating rhythms in insects. In *Drosophila*, null mutants of the clock gene *period* (*per*) lost the circadian rhythm in mating activity (Sakai and Ishida 2001). Similar effects have been described for gene *timeless* (*tim*) in *Drosophila* and for gene *per* in *Grillus bimaculatus* De Geer, 1773 (Sehgal et al. 1994, Moriyama et al. 2008). The analysis of mating activity in transformant lines carrying *per* transcription units derived from *Drosophila melanogaster* Meigen, 1930 or *Drosophila pseudoobscura* Frolova & Astaurov, 1929, showed that *per* controls species-specific mating rhythms, at least in flies (Tauber et al. 2003).

It may be suggested that differences in the rhythm of mating activity in two forms of *C. pipiens* resulted from the variations in circadian clocks genes. To test this hypothesis, we selected the genes *per* and *tim*. The aim of our work was to study variable nucleotide sequences in these genes, and to estimate the possible evolutionary significance of the detected variations.

Methods

The larvae of mosquitoes of both intraspecific forms were collected mostly in August 2006 in Volgograd City and nearby areas. The sampling sites, methods of larvae collection and rearing in lab, and methods of evaluating autogenity have been described earlier (Fedorova and Shaikevich 2013). The DNA of mosquitoes collected in the underground sampling sites in Nizhny Novgorod, Moscow and St Petersburg, as well as in aboveground sampling site Iksha, Moscow region, was used to analyze the diversity of the first exon of the gene *tim*. The methods of mosquito sampling at these sites have been described earlier (Vinogradova and Shaikevich 2007).

DNA isolation and analysis

The DNA was isolated using the kit DIAtom™ DNA Prep (Isogen Russia). Each of the amplification reactions used 0.1 μg of the total DNA. The polymerase chain reaction (PCR) was run on the thermocycler GeneAmpR PCR System 2700 (Applied Biosystems USA), with amplification Encyclo PCR kit (Evrogen Russia), following the manufacturer's instructions. For PCR and sequencing of amplification products, specific primers were constructed which were complementary to the conserved sequences of exons in the published sequences of the genes *period* and *timeless* from the total genome of a similar species *C. quinquefasciatus* (Vector Base Gene ID CPIJ007193 and CPIJ007082, respectively) (Arensburger et al. 2010). When the first sequences were obtained, the primers were constructed basing on DNA sequences of *C. pipiens*. The PCR conditions were adjusted using the program Oligo6 (http://www.oligo.net/): primary denaturing 95°C - 5 min; 35 cycles at 95°C - 30 s, Tm (for each primer pair) - 1 min, 72°C - 1,5 min; final synthesis at 72°C for 7 min. Primer sequences and annealing temperatures for the PCR are shown in Table 1. Higher temperature was used if two primers in the pair had different annealing temperatures. Negative control was run for all amplification reactions. The DNA of introns was analysed by direct sequencing of amplicons without cloning. Amplified fragments of the genes *per* and *tim* were purified from the gel using QIAquick Gel Extraction kit (Qiagen USA). The fragments were cloned using the kit pGEM-T Easy Vector Systems (Promega USA); the DNA of the three clones for each individual mosquito was sequenced using the equipment ABI PRISM 310 and the BigDye Termination kit (Applied Biosystems USA), according to the manufacturer's instructions and deposited to GenBank under accession numbers: KU133680-KU133745. The sequences of separate exons of each clone were combined into a single sequence. Nine combined sequences from individual *C. pipiens* f. "pipiens" and nine combined sequences from individual *C. pipiens* f. "molestus" were investigated for each of the two genes studied, *per* and *tim*. Extended study of the coding sequences of exon 1 of the gene *tim* in two forms of *C. pipiens* was performed using the DNA from the 26 individual mosquitoes *C. pipiens* f. "molestus" and 17 individual mosquitoes *C. pipiens* f. "pipiens". 21 new different haplotypes are submitted to GenBank (KU997646 - KU997666).

Data analysis

The DNA sequences were translated into amino acids sequences using ExPASy software (Swiss Institute of Bioinformatics), and compared with amino acids sequences of *C. quinquefasciatus* (Arensburger et al. 2010) and *C. pipiens* from the USA (Meuti et al. 2015) using programs MAFFT (http://mafft.cbrc.jp/alignment/server/) and MEGA6 (Tamura et al. 2013). Evolutionary analysis was run using MEGA6. Maximum Composite Likelihood model (Tamura et al. 2004) and Kimura 2-parameter model (Kimura 1980) were used to describe the nucleotide substitution pattern. Tables below show

Table 1. Primers constructed to study the genes *per* and *tim*.

primer	sequence	Tm (°C)	region
PerF2	5'-AGTTCCAAATCGCGCCACAG-3'	54	*per* exon 2
PerR2	5'-TTGGGTTTGCTCGCTTCGTTC-3'	54	*per* exon 2
PerF3	5'-ACAATGCATAGCCAACCGCAAG-3'	55	*per* exon 3
PerR3	5'-GTTCGTCCCTTGACCATGATC-3'	54	*per* exon 3
PerF4	5'-AACGGCTGTTATCTCGTACTG-3'	52	*per* exon 4
PerR4	5'-GCATCGCGTGGTACATCATCG-3'	56	*per* exon 4
TimF1	5'-AATGGTTGCTAGCGAATCCG-3'	52	*tim* exon1
TimR1	5'-AGTAGAGTTCTCGACACCCG-3'	54	*tim* exon1
TimF5	5'-GATTGGTCGGATTTGATTGAG-3'	50	*tim* exon5
TimR5	5'-GTATGTCATCAACCGCCTTG-3'	52	*tim* exon5
TimF5-1	5'-GGAAACCAGCAAAAGACTCG-3'	52	*tim* intron5-6, exon6, intron6-7
TimR7	5'-TACGAGAGCACGTTGAACTG-3'	52	*tim* intron5-6, exon6, intron6-7
TimF7	5'-ACATACTGTACAACATTGCCCTG-3'	53	*tim* intron7-8
TimR8	5'-TCAGGTCGAACTTGATGATG-3'	50	*tim* intron7-8
TimF9	5'-GCTGCGGCCGAAAGCGCCAG-3'	60	*tim* intron9-10
TimR10	5'-ATTTCCATCGCTCGTGTGCTG-3'	54	*tim* intron9-10

the data obtained using the Maximum Composite Likelihood model. The Kimura 2-parameter model produced somewhat higher estimates. The optimal model describing evolutionary patterns was found using the option 'Find best DNA/Protein substitution model' in MEGA6. For our data, the Jones-Taylor-Thornton (JTT) model (Jones et al. 1992) showed the lowest BIC (Bayesian Information Criterion) scores for amino acid sequences and was selected to describe the amino acids substitution pattern. For estimating polymorphism within each group and evolutionary divergence between each group the number of base substitutions per site from averaging over all sequence pairs was calculated, all positions containing gaps and missing data were eliminated. Codon positions included were 1st+2nd+3rd+Noncoding. For the estimation of Maximum Likelihood Estimate of Transition/Transversion Bias (R) substitution pattern and rates were estimated under the Kimura 2-parameter model.

Phylogeny analysis was run in MEGA6. The evolutionary history was inferred using the Neighbor-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. All ambiguous positions were removed for each sequence pair.

Natural selection and the probability of rejecting the null hypothesis of strict-neutrality (dN = dS) was evaluated using MEGA6. For these purposes was used a codon-based Z-test (MEGA6). For a pair of sequences, this is done by first estimating the number of synonymous substitutions per synonymous site (dS) and the number of nonsynonymous substitutions per nonsynonymous site (dN), and their variances: Var(dS) and Var(dN), respectively. With this information, we tested the null hypothesis that there is no impact of selection (dN = dS) and the probability (P) of rejecting the null hypothesis of strict-neutrality. Also was tested an alternative hypothesis of pu-

rifying selection (dN < dS) and the probability of rejecting the null hypothesis of strict-neutrality in favor of the alternative hypothesis using a codon-based Z-test (MEGA6). Values of *P* determine statistical significance in a hypothesis test. A low *P* value suggests that sample provides enough evidence for the rejecting of the null hypothesis for the entire population. Values of *P* less than 0.05 are considered significant at the 5% level. The variance of the difference was computed using the analytical method (Kimura 1980). All ambiguous positions were removed for each sequence pair.

Results

The gene *period* (*per*) in two forms of *C. pipiens*

The structure of the gene *per* was studied in three individual *C. pipiens* f. "molestus" and in three individual *C. pipiens* f. "pipiens". Coding sequences of the three exons of the gene *per* were analysed: exon 2, 333 bp, exon 3, 738 bp, and exon 4, 1229 bp. In total, the 18 compared sequences spanned each 2300 bp (Suppl. material 1).

In the exon 2 of the gene *per* (333 bp) 11 variable sites were found; six of these substitutions resulted in amino acid substitutions in both forms of *C. pipiens* (Fig. 1). The exon 3 (738 bp) had 12 variable nucleotide sites, resulting in three amino acid substitutions (Fig. 1). The exon 4 (1229 bp) had 27 variable nucleotide sites, resulting in four amino acid substitutions (Fig. 1). In total, the nucleotide sequence of the three exons of the gene *per* for the both intraspecific forms had 50 (2.2%) variable nucleotide sites and 13 (1.7%) polymorphic amino acid sites, 48 nucleotide sites being parsimony-informative. The estimated Transition/Transversion bias (R) is 2.83. The DNA polymorphism of the gene *per* among individuals of the *C. pipiens* f. "pipiens" (0.003) and of the *C. pipiens* f. "molestus" (0.002) were both low, variability of the amino acid sequences also was low (Table 2). The genetic distances between two forms of *C. pipiens* from Volgograd were 0.010 based on nucleotide sequences and 0.011 based on amino acid sequences of the gene *per* (Table 2).

Comparison of the gene *per* for transatlantic *C. pipiens*

The obtained sequences of the gene *per* of *C. pipiens* from Volgograd and *C. pipiens* f. "pipiens" from the USA (GenBank acc. number KM355980) using BLAST software were compared. The identity of nucleotide sequences of *C. pipiens* f. "pipiens" mosquitoes from different continents is 98-99%; 4-16 amino acid substitutions were detected. Pairwise comparison showed that *C. pipiens* f. "pipiens" from the USA is slightly different from the Volgograd *C. pipiens* f. "pipiens" (0.008) and from *C. pipiens* f. "molestus" (0.013); these values are comparable with the differences between the studied *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus" (Table 2).

```
[                                  11  111111111223  5555556677777]
[                  2455566778999900  123345558581  1234793646666]
[                  5115658158012425  704922470845  6538609040124]
#molestus2         VMVSGAAMECASAGQN  MSKQVPVASSSK  DTSDGLNESTPTP
#molestus1         ................  ............  ..C.R..D.....
#molestus3         A...............  ........T...  .......D.....
#pipiens1          ....E....TTTT...  IN......T...  .........I...
#pipiens2          .........TTTT...  I...........  .........I...
#pipiens3          .........TTTT...  I...........  .........I...
#pipiensUSA_KM355980  .........TTTP..D  ............  E.......NI...
#quinq_CPIJ007193     .TAC.STLDT..TVP.  ..EAAMAP.TIR  .I.G.QHD..SI.
```

Figure 1. Variable amino acid sites of the gene *period* in *C. pipiens*. *C. pipiens* from the USA (KM355980) and *C. quinquefasciatus* (CPIJ007193) are taken for the comparison. Exons 2 (sites 1-111), 3 (113-358), and 4 (360-768) are separated with blank columns. Positions of the variable sites relative to combined sequences as presented in Suppl. material 1 shown on the top.

Table 2. Estimates of Evolutionary Divergence over *per* and *tim* sequence pairs between *Culex pipiens* complex members.

	AA\NA	gene *period*				gene *timeless*			
		1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1	molestus		**0.010**	**0.010**	**0.027**		**0.012**	**0.025**	**0.028**
2	pipiens	0.011		**0.008**	**0.029**	0,008		**0.031**	**0.030**
3	pipiensUSA	0.013	0.008		**0.027**	0.018	0.018		**0.009**
4	quin	0.036	0.036	0.041		0.018	0.020	0.004	

In upper right section in bold: the number of nucleotide base substitutions (NA) per site from averaging over all sequence pairs between groups is shown. All results are based on the pairwise analysis of 20 sequences. There were a total of 2300 positions of *per* gene and 1560 positions of *tim* gene in the final dataset. In lower left section: the number of amino acid substitutions (AA) per site from averaging over all sequence pairs between groups are shown. The analysis involved 20 amino acid sequences. A total of 766 positions of the gene *per* and 520 positions of the gene *tim* were analysed as the final dataset.

Comparison of the gene *per* in *C. pipiens* and *C. quinquefasciatus*

The identity of the nucleotide sequences of the gene *per* for the two species was 97%. Comparison of the DNA from both forms of *C. pipiens* and *C. quinquefasciatus* (CPIJ007193) revealed 113-116 (4.9–5.5%) variable nucleotide sites (Suppl. material 1): 37 nucleotide substitutions were non-synonymous, resulting in amino acid substi-

tutions (Fig. 1). 64 nucleotide substitutions and 24 amino acid substitutions are specific for *C. quinquefasciatus,* with nine substitutions in each of the exons 2 and 3, and six in exon 4 (Fig. 1). The mean genetic divergence between *C. pipiens* and *C. quinquefasciatus* is 0.03 by both DNA and amino acid sequences. The difference between *C. pipiens* and *C. quinquefasciatus* is three times higher than the difference between the two forms of *C. pipiens* (Table 2).

Gene *timeless* (*tim*) in the two forms of *C. pipiens*

Using the DNA from the same three individual mosquitoes *C. pipiens* f. "molestus" and three individual mosquitoes *C. pipiens* f. "pipiens", the three longest coding sequences of the gene *tim* were studied: exon 1 (1037 bp), exon 5 (376-379 bp), and exon 6 (145 bp). In total, the 18 compared sequences each spanned 1557-1560 bp. (Suppl. material 2).

In exon 1 of the gene *tim* (1037 bp) 15 variable nucleotide sites and five variable amino acid sites were found, four of them showing variations only for the *C. pipiens* f. "pipiens" and were not found in *C. pipiens* f. "molestus" (Fig. 2). In exon 5 (379 bp) five DNA substitutions were found, and in exon 6 (145 bp) there were three variable nucleotide sites; all substitutions in the exons 5 and 6 were synonymous, resulting in similar amino acid sequences for *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus" (Suppl. material 2, Fig. 2). The estimated Transition/Transversion bias (R) is 2.79.

Comparing the nucleotide sequences of the gene *tim* for the specimens of *C. pipiens* f. "molestus" one variable DNA site was found, the detected nucleotide substitution does not result in amino acid substitution. The aligned DNA sequences of the *C. pipiens* f. "pipiens" had 11 variable sites, one mutation resulting in amino acid substitution (Fig. 2). The DNA polymorphism of the gene *tim* among specimens of the *C. pipiens* f. "pipiens" (0.0038) was higher than for *C. pipiens* f. "molestus" (0.0004), and variability of the amino acid sequences was 0.001 and 0.000, respectively. Comparing the total sequence of the three exons of the gene *tim,* between *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus" 23 (1.5%) variable nucleotide sites were found (all 23 sites were parsimony-informative) and six (0.4%) polymorphic amino acid sites. Genetic distance between the two forms was 0.012 for DNA sequences and 0.008 for amino acid sequences (Table 2).

Comparison of the gene *tim* for the transatlantic *C. pipiens*

The obtained sequences of the gene *tim* for *C. pipiens* from Volgograd and *C. pipiens* f. "pipiens" from the USA (KM355979) were compared using BLAST software. Identity of nucleotide sequences for mosquitoes of *C. pipiens* f. "pipiens" from different continents is 96-97%. We found 7-12 amino acid substitutions. Unexpectedly, we found a 60-bp deletion within the coding sequence of exon 1 in *C. pipiens* f. "pipiens" from the USA (KM355979), positions 263-282 in Fig. 2. No similar deletion was found in either of the studied *C. pipiens* forms from Volgograd and no similar deletions were

```
[                          1122222222222222222222233333    3]
[                        923246666666777777777888900223    7]
[                        196063456789012345678901240202O   5]
#molestus2               YVNCDGDSLTLGWDHQCTQTTLLIPELTTHT    E
#molestus1               ..............................     .
#molestus3               ..............................     .
#pipiens1                ....E.......................SQI    .
#pipiens2                ....E..............P.......SQI    .
#pipiens3                ....E..............P.......SQI    .
#pipiensUSA_KM355979     FIDF.--------------------DQSSQ.    .
#quinq_CPIJ007082        FI.F..................P..DQSSQ.    Q
```

Figure 2. Variations of amino acid sites in the gene *tim* from *C. pipiens*. *C.pipiens* from the USA (KM355979), and *C.quinquefasciatus* (CPIJ007082) are taken for the comparison. Dashes show deletion in exon 1 in *C. pipiens* from the USA (KM355979). Exon 1 (sites 1-345) and exon 5 (sites 347-472) are separated by blank columns. Positions of the variable sites in combined *tim* sequences shown on the top.

found in *C. quinquefasciatus* (CPIJ007082). The genetic distance between *C. pipiens* f. "pipiens" from Volgograd and *C. pipiens* f. "pipiens" from the USA (KM355979) is 0.025, two times higher than the distance between both forms of *C. pipiens* in Volgograd 0.012 (Table 2).

Comparison of the gene *tim* from *C. pipiens* and *C. quinquefasciatus*

Comparison of DNA sequences of exons 1, 5 and 6 of the gene *tim* between representatives of the two species, *C. pipiens* and *C. quinquefasciatus,* revealed 50 variable sites (see Suppl. material 2), which result in 8-14 amino acid substitutions, eight of which are found only in *C. quinquefasciatus* (Fig. 2). The genetic distance between the species is 0.029 in the DNA sequences and 0.019 in amino acid sequences (Table 2). A striking similarity should be noted for the gene *tim* from *C. quinquefasciatus* and *C. pipiens* f. "pipiens" from the USA (KM355979). Their amino acid sequences have only three variable sites: 136, 280, and 375 (Fig. 2), their DNA sequences differ in 13 single-nucleotide substitutions and one deletion. The genetic distance for the gene *tim* between *C. quinquefasciatus* and *C. pipiens* f. "pipiens" from the USA (KM355979) is 0.009 based on DNA sequences and 0.004 based on amino acid sequences, lower that the distance between the two forms of *C. pipiens* in Volgograd (Table 2).

Extended study of exon 1 of the gene *tim* in two forms of *C. pipiens*

Our results showed that exon 1 of the gene *tim* in *C. pipiens* f. "molestus" differ from that of in *C. pipiens* f. "pipiens" (Fig. 2). Contrary to the gene *per,* no shared polymorphisms were found in amino acid sequences of gene *tim* between two forms (Figs 1, 2). To confirm these findings we studied the structure of exon 1 (1037 bp) of the gene *tim* in 23 specimens of *C. pipiens* f. "molestus" and 14 specimens of *C. pipiens* f. "pipiens" in addition to 6 samples of gene *tim* described above. In total, 43 samples were examined.

The obtained nucleotide sequences showed overlapping peaks in one or more sites for 6 individuals. Four of them were identified as *C. pipiens* f. "molestus" and two as *C. pipiens* f. "pipiens". Exon 1 of the gene *tim* of these six samples was studied by cloning and the DNA of the five clones for each specimen was sequenced. In total 79 sequences were obtained for comparative analysis (Suppl. material 3). In five specimens one allele was identical to *C. pipiens* f. "pipiens" and other one was identical to *C. pipiens* f. "molestus", i. e. these mosquitoes represented hybrids. In one *C. pipiens* f. "molestus" (NN23) the alleles differed by two nucleotide substitutions in 3'end. All hybrids were collected in Volgograd, where both forms develop in the same pools in summer.

49 variable nucleotide sites and 23 distinct haplotypes were found in exon 1 of the gene *tim* (Fig. 3). *C. pipiens* f. "pipiens" showed 19 haplotypes. Four haplotypes were obtained in *C. pipiens* f. "molestus" (H1-H4). Haplotypes H1 and H2 detected in *C. pipiens* f. "molestus" from geographically remote locations (Volgograd, Nizhny Novgorod, Moscow and S.-Petersburg) differed by only one synonymous nucleotide substitution A-G at position 653 in Exon 1 of the gene *tim* (Fig. 3). H3 and H4 were detected only in two individuals: H3 combined with H1 (*C. pipiens* f. "molestus") in NN23 and H4 in combination with H11 (*C. pipiens* f. "pipiens") in V219 (Suppl. material 3). Amino acid sequences of *C. pipiens* f. "molestus" with H1 and H2 haplotypes differed from *C. pipiens* f. "pipiens" by two substitutions Serine (Ser)-Threonine (Thr) and *Glutamine* (Gln)-Histidine (His). Additional substitutions were detected in two specimens with H3 and H4 haplotypes namely the T968A and T968G substitutions in DNA sequences which resulted in Gln in amino acid sequence (Fig. 3, Suppl. material 3). In total, two variations of amino acid sequences were found in *C. pipiens* f. "molestus" and 8 in *C. pipiens* f. "pipiens" (Fig. 3).

The DNA polymorphism of the exon 1 of gene *tim* among specimens of the *C. pipiens* f. "pipiens" (0.007) was ten times higher than for the *C. pipiens* f. "molestus" (0.0006), and variability of the amino acid sequences was 0.0053 and 0.0001, respectively. Genetic distance between the two forms was 0.011 for DNA sequences and 0.009 for amino acid sequences. Genetic distance between *C. pipiens* of both forms and *C. quinquefasciatus* was 0.029 for DNA and 0.02 for amino acid sequences (Suppl. material 3). The DNA polymorphism, as well as genetic distances between the two forms in extended study of the exon 1 are very close to the results obtained for the three exons of the gene *tim* (see above) (Table 2).

```
[                                  1]
[         1112223 3445555667 77899990]  122333
[      5794460674 8671245572 47366690]  547223
[      3764940991 1178434315 00402814]  968020
#H1    CTATTCCCTT CGCTTTCAAG TAAAGTCT   PDTTHT
#H2    .......... .......G.. ........   ......
#H3    .......... .......... .....A.A   ....Q.
#H4    ..C...T... T..C.C.G.. .....G..   ....Q.
#H5    .A.CCTT..C .......G.. A..T.GT.   .E.SQI
#H6    .A.CCTTG.C .A.....G.. A.CTAGT.   .EPSQI
#H7    .A..C.TG.C .A.....G.. A.CTAGT.   .EPSQI
#H8    .A.C.TTG.C .A.....G.. A.CTAGT.   .EPSQI
#H9    .A.CCTTG.C .A.....G.. A.CTTGT.   .EPSQI
#H10   .A...TT... ..T.C..G.A .TCT.G..   S.PSQ.
#H11   .A.CCTT..C ..T...TG.A ...T.G..   S..SQ.
#H12   .A.CCTT... ..T.C..G.A .T.T.G..   S..SQ.
#H13   .A...TT.C. ..T.C..G.A .T.T.G..   S..SQ.
#H14   .A...TT..C ..T.C..G.A .T.TAGT.   S..SQI
#H15   ...CCTT..C .A.....G.. A.-T.GT.   .E-SQI
#H16   .....TT... ....C..G.. ...TTG..   ...SQ.
#H17   .A...TT.CC .......G.A ...T.G..   ...SQ.
#H18   .A...TT... ....C..GGA ...T.G..   ...SQ.
#H19   T..CCT...C .......G.A .T.TAG..   ...SQ.
#H20   .A...TT... .......G.. .T.T.G..   ...SQ.
#H21   .A.CCTT..C .......G.. ...T.GT.   ...SQI
#H22   .A.CCTT..C .A..C..G.. ..CT.G..   ..PSQ.
#H23   .....TT... .......GGA ..CT.G..   ..PSQ.
```

Figure 3. DNA haplotypes and variable amino acid positions in the exon 1 of the gene *tim* from *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus". Haplotypes numbers and variable nucleotide sites are shown on the left. Variable amino acid sites are shown on the right. Only variable haplotypes are shown, all 79 sequences are presented in Suppl. material 3. Positions of the variable sites shown on the top. Dash show deletion of 12 nucleotides (sites 831-842) in exon 1 in *C.pipiens* f. "pipiens" from Moscow region.

Variation in non-coding regions of the gene *tim*

The sequences of some non-coding regions were analysed, expecting to find differences not only in coding DNA structure but also in intron size between *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus". The primers were constructed for the conserved sites of the exons using the obtained sequences, and by homology with the gene *tim* from *C. quinquefasciatus* (CPIJ007082). The sequences of the introns 1-2 (7158 bp in length) and 10-11 (5189 bp), being too long for efficient PCR and sequencing and containing numerous repeats were not analysed. As for the other introns, sequencing of the PCR products showed no variability between two intraspecific forms in intron between exons 5 and 6 (59 bp). In intron 6-7 (61 bp) three variable sites and in intron 7-8 (160 bp) six variable sites were found. Studied introns showed no mutations common with either of the two *C. pipiens* forms. In the intron 9-10 (167 bp) seven variable sites were found, six of which differed between the two forms (Suppl. material 4). The length of all amplified intron sequences was identical for *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus".

Phylogenetic analysis

Phylogenetic dendrograms were constructed applying the Neighbor-Joining method to amino acid sequences of the three coding regions of genes *per* and *tim*, *C. quinquefasciatus* was used as an out-group. *C. quinquefasciatus* and *C. pipiens* form two well differentiated clusters. Basing on similarity of the gene *per* the individuals of the *C. pipiens* f. "pipiens" group together and form a joint cluster with a bootstrap coefficient of 97. The studied specimens of the *C. pipiens* f. "molestus" had more polymorphic amino acid sequences, but also are grouped into one cluster with bootstrap coefficient of 63 (Fig. 4). Based on the similarity of the gene *tim*, *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus" group into separate clusters with a bootstrap coefficient of 88 (Fig. 4B).

On the dendrogram for the exon 1 of the gene *tim*, constructed using the results of our extended study, most specimens of the *C. pipiens* f. "molestus" form separate clusters with a bootstrap coefficient of 96. A separate subcluster is formed by sequences of the hybrid V219 clones with haplotype H4. The studied specimens of the *C. pipiens* f. "pipiens" have polymorphic DNA sequences (Suppl. material 3). The dendrogram basing on amino acid sequences shows similar configuration.

Evolutionary analysis

One way to test whether natural selection is operating on a gene is to compare the relative abundance of synonymous and nonsynonymous substitutions within the gene sequences (Tamura et al. 2013). Analysing evolution of the nucleotide sequences, the

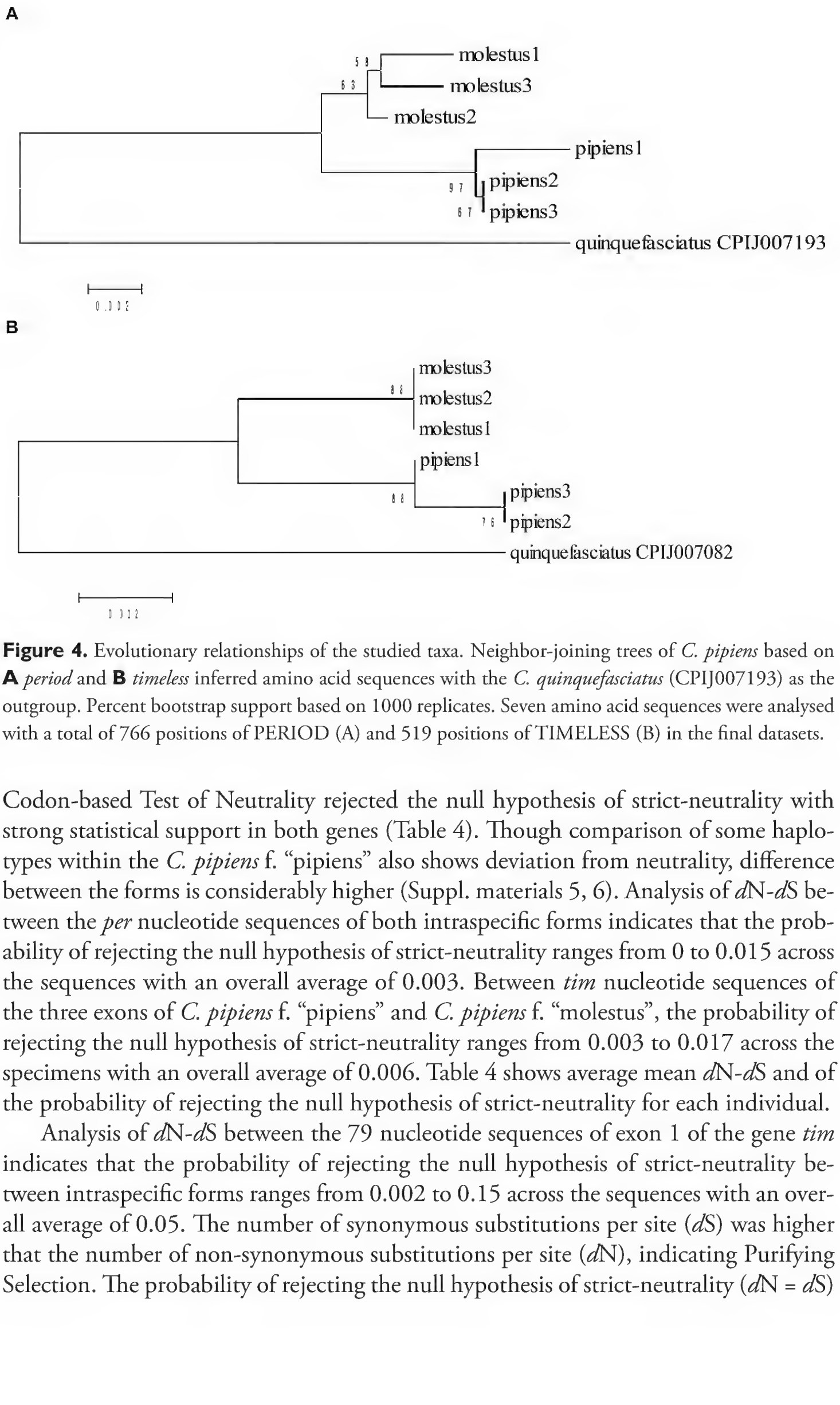

Figure 4. Evolutionary relationships of the studied taxa. Neighbor-joining trees of *C. pipiens* based on **A** *period* and **B** *timeless* inferred amino acid sequences with the *C. quinquefasciatus* (CPIJ007193) as the outgroup. Percent bootstrap support based on 1000 replicates. Seven amino acid sequences were analysed with a total of 766 positions of PERIOD (A) and 519 positions of TIMELESS (B) in the final datasets.

Codon-based Test of Neutrality rejected the null hypothesis of strict-neutrality with strong statistical support in both genes (Table 4). Though comparison of some haplotypes within the *C. pipiens* f. "pipiens" also shows deviation from neutrality, difference between the forms is considerably higher (Suppl. materials 5, 6). Analysis of *d*N-*d*S between the *per* nucleotide sequences of both intraspecific forms indicates that the probability of rejecting the null hypothesis of strict-neutrality ranges from 0 to 0.015 across the sequences with an overall average of 0.003. Between *tim* nucleotide sequences of the three exons of *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus", the probability of rejecting the null hypothesis of strict-neutrality ranges from 0.003 to 0.017 across the specimens with an overall average of 0.006. Table 4 shows average mean *d*N-*d*S and of the probability of rejecting the null hypothesis of strict-neutrality for each individual.

Analysis of *d*N-*d*S between the 79 nucleotide sequences of exon 1 of the gene *tim* indicates that the probability of rejecting the null hypothesis of strict-neutrality between intraspecific forms ranges from 0.002 to 0.15 across the sequences with an overall average of 0.05. The number of synonymous substitutions per site (*d*S) was higher that the number of non-synonymous substitutions per site (*d*N), indicating Purifying Selection. The probability of rejecting the null hypothesis of strict-neutrality (*d*N = *d*S)

Table 3. Comparison of exons and introns variability between *C. pipiens* f. "pipiens" and f. "molestus" from Russia.

Gene	Locus	size (bp)	Variable DNA sites	Differentiating DNA sites	Variable AA sites	Differentiating AA sites
Per	exon2	333	11	6	6	4
	exon3	738	12	2	3	1
	exon4	1229	27	9	4	1
Tim	exon1	1037	49	2	10	1
	exon5	376–379	5	1	0	0
	exon6	145	3	0	0	0
	intron5-6	59	0	0	-	-
	intron6-7	61	3	0	-	-
	intron7-8	160	6	0	-	-
	intron9-10	167	7	6	-	-

AA - amino acid

in favor of the alternative hypothesis of Purifying Selection ($dN < dS$) ranges for the *tim* nucleotide sequences of *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus" from 0.001 to 0.10 with an overall average of 0.025 (Suppl. material 7).

Discussion

For the first time the genetic structure of the circadian rhythm genes (*per* and *tim*) were analysed for mosquitoes *C. pipiens* f. "molestus". Our results have shown that DNA variation in individuals of *C. pipiens* f. "molestus" is smaller than in individuals of *C. pipiens* f. "pipiens". Extended study of exon 1 of the gene *tim* revealed 4 DNA haplotypes in *C. pipiens* f. "molestus" and 19 haplotypes in *C. pipiens* f. "pipiens". Decrease in DNA variability for the underground mosquitoes of *C. pipiens* f. "molestus" was also reported earlier in our study of mitochondrial DNA (Shaikevich and Zakharov 2010).

In coding sequences of both genes *per* and *tim*, variations between physiologically different forms of *C. pipiens* were found (Table 3). In the gene *per* we found nine polymorphisms shared between the two forms and four fixed differences between the two forms, taking into account *C. pipiens* f. "pipiens" from N America (Fig. 1). The gene *tim* had one shared amino acid polymorphisms and one fixed difference between the forms (Fig. 3). Higher variation of the gene *per* is also revealed by comparison of *C. pipiens* and *C. quinquefasciatus*: basing on the amino acid sequences, the genetic distances between the species are higher for the gene *per* (0.036) that for the gene of *tim* (0.02).

C. pipiens f. "pipiens" from N America clusters with *C. pipiens* f. "pipiens" from Volgograd basing on comparison of the gene *per* and with *C. quinquefasciatus* based on comparison of the gene *tim*. It remains unknown whether this is common for all American *C. pipiens* f. "pipiens", shown using microsatellite analysis to differ from the

Table 4. Codon-based Test of Neutrality between *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus".

specimen		gene *period*						gene *timeless*					
		1	2	3	4	5	6	1	2	3	4	5	6
1	pipiens1		-1.950	-2.829	-3.130	-3.682	-3.309		-2.307	-2.009	-2.830	-2.779	-2.679
2	pipiens2	0.108		-2.904	-3.326	-3.865	-3.349	0.034		-1.856	-2.906	-2.859	-2.766
3	pipiens3	0.010	0.008		-3.775	-4.284	-3.788	0.089	0.0663		-2.945	-2.899	-2.805
4	molestus1	0.001	0.003	0.000		-1.534	-0.586	0.0047	0.0023	0.005		-0.333	-0.998
5	molestus2	0.000	0.000	0.000	0.154		-1.427	0.0056	0.0027	0.0058	0.774		-0.665
6	molestus3	0.002	0.003	0.000	0.502	0.15		0.0073	0.0034	0.0075	0.320	0.547	

The test statistic (dN - dS) is shown above the diagonal. The probability of rejecting the null hypothesis of strict-neutrality (dN = dS) is shown. Values of P less than 0.01 are considered significant at the 1% level. There was a total of 766 positions of gene *per* and of 519 positions of gene *tim* in the final dataset. Evolutionary analyses were conducted in MEGA6.

European *C. pipiens* f. "pipiens" (Fonseca et al. 2004), or if it is a specific feature of the laboratory line, used to study the genes on circadian rhythm (Meuti et al. 2015).

Genetic structure of the studied genes is polymorphic. However, the revealed substitutions in nucleotide sequences and especially in protein sequences grouped the individuals of the two forms into distinct clusters with high significance, a longer genetic distance separating the cluster of *C. pipiens* from *C. quinquefasciatus*. Although the two studied genes differed in variability, the results of analysis of the gene *per*, as well as the gene *tim*, show that the difference between *C. pipiens* and *C. quinquefasciatus* are 2.5–3 times higher than the difference between the forms of *C. pipiens*. The genetic distances again confirm the order of evolutionary events in the *C. pipiens* complex: the divergence of the form *C. pipiens* f. "molestus" from *C. pipiens* occurred considerably later than the divergence of *C. pipiens* and *C. quinquefasciatus* (Barr 1967, Fonseca et al. 2004, Shaikevich and Zakharov 2014).

The non-coding genome sequences are considered to be highly variable. These sequences are often used to search for the markers to differentiate closely related organisms by size of the PCR products. For example, variation in spacers of the ribosomal genes cluster is a base for identification of some mosquito species of the genus *Anopheles* (Nicolescu et al. 2004, Gordeev et al. 2004). In sequences of three *tim* introns no significant difference was found between the forms. For *Aedes albopictus* Skuse, 1894, also no significant difference in the introns of the gene *tim* was reported (Summa et al. 2012).

The Test of Neutrality rejects the null hypothesis of strict-neutrality at $P < 5\%$ level and imply that both *per* and *tim* loci evolve under strong selective constraint during the divergence of intraspecific forms. Our results suggest that natural selection favored the fixed mutations and the decreased diversity of the genes *per* and *tim* in mosquitoes *C. pipiens* f. "molestus" compared with the *C. pipiens* f. "pipiens", probably preserving adaptive features of the form "molestus". Well-documented data have been reported showing that new native mutations sometimes are rapidly spreading in a population and that polymorphism in one locus may provide adaptive variations in behavioral and morphological phenotypes of the insects in nature (Tauber et al. 2007). The genes

involved in circadian rhythms are proved to coordinate seasonal responses, e.g. they initiate the reproductive diapause; malfunctioning of the genes *per* and *tim* was shown to interrupt diapausing of the *C. pipiens* females (Meuti et al. 2015). We can assume that mutations found in *per* and especially in *tim* genes are related with functioning of the circadian rhythm proteins and contributed to divergence of the forms of *C. pipiens.* The studied genes are promising candidates to evaluate the genetic basis of different behaviors of the two ecological forms within one subspecies. Further studies of the circadian rhythm genes in mosquitoes of the *Culex pipiens* complex would help to test this assumption.

Conclusions

Nucleotide sequences of the circadian rhythm genes were studied for the first time in mosquitoes *C. pipiens* f. "molestus" and compared with those for *C. pipiens* f. "pipiens" and *C. quinquefasciatus.* These results show that intraspecies variability is higher for the gene *per* than for the gene *tim.* Revealed substitutions in nucleotide sequences and especially in protein sequences grouped the individuals of the two ecological forms of *C. pipiens* into distinct clusters with high significance. The results suggest that natural selection favored the fixed mutations and the decreased diversity of the genes *per* and *tim* in mosquitoes of the *C. pipiens* f. "molestus" compared with the *C. pipiens* f. "pipiens". The detected fixed amino acid substitutions may appear essential for functioning of the circadian rhythm proteins in *C. pipiens*, and may be related with adaptations of the taxa within the group *C. pipiens*. Moreover, under natural selection mutations in the key genes of circadian pattern may provide some advantage to the underground *C. pipiens* f. "molestus". The studied genome regions may be considered as promising molecular-genetic markers for identification, population and phylogenetic analysis of similar species and forms of the *C. pipiens* complex.

Acknowledgements

This work was supported by the Russian Foundation of Fundamental Research, grants N 14-04-0112914, N 16-04-00091 and the European Commission in the framework of FP7-261391 EuroWestNile research project.

References

An X, Wilkes K, Bastian Y, Morrow JL, Frommer M, Raphael KA (2002) The period gene in two species of tephritid fruit fly differentiated by mating behaviour. Insect Molecular Biology 11: 419–430. doi: 10.1046/j.1365-2583.2002.00351.x

An X, Tebo M, Song S, Frommer M, Raphael KA (2004) The cryptochrome (cry) gene and a mating isolation mechanism in Tephritid fruit flies. Genetics 168: 2025–2036. doi: 10.1534/genetics.104.028399

Arensburger P, Megy K, Waterhouse RM, Abrudan J, Amedeo P, Antelo B et al. (2010) Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. Science 330: 86–88. doi: 10.1126/science.1191864

Bahnck CM, Fonseca DM (2006) Rapid assay to identify the two genetic forms of *Culex* (*Culex*) *pipiens* L. (Diptera: Culicidae) and hybrid populations. American Journal of Tropical Medicine and Hygiene 75: 251–255.

Barr AR (1967) Occurrence and distribution of the *Culex pipiens* complex. Bulletin of World Health Organization 37: 293–296.

Bertossa RC, van Dijk J, Diao W, Saunders D, Beukeboom LW et al. (2013) Circadian Rhythms Differ between Sexes and Closely Related Species of Nasonia Wasps. PLoS ONE 8(3): e60167. doi: 10.1371/journal.pone.0060167

Byrne K, Nichols RA (1999) *Culex pipiens* in London underground tunnels: differentiation between surface and subterranean populations. Heredity 82: 7–15. doi: 10.1038/sj.hdy.6884120

Fedorova MV, Shaikevich EV (2013) Role of the mosquitoes *Culex pipienis* f. pipiens and *Cx. pipiens* f. molestus (Diptera, Culicidae) in the spread of West Nile virus in the south of Russia. Meditsinskaia parazitologiia i parazitarnye bolezni 3: 36–39. PubMed ID 25850309. [In Russian]

Fergus DJ, Shaw KL (2013) Circadian rhythms and period expression in the Hawaiian cricket genus Laupala. Behavior Genetics 43(3): 241–253. doi: 10.1007/s10519-012-9576-4

Fonseca DM, Keyghobadi N, Malcolm CA, Mehmet C, Schaffner F, Mogi M, Fleischer RC, Wilkerson RC (2004) Emerging vectors in the *Culex pipiens* complex. Science 303: 1535–1538. doi: 10.1126/science.1094247

Fuchikawa T, Sanada S, Nishio R, Matsumoto A, Matsuyama T, Yamagishi M, Tomioka K, Tanimura T, Miyatake T (2009) The clock gene cryptochrome of Bactrocera cucurbitae (Diptera: Tephritidae) in strains with different mating times. Heredity (Edinburgh) 104(4): 387–392. doi: 10.1038/hdy.2009.167

Fyodorova MV, Serbeniouk SA (1999) Evaluation of male reproductive success trough body size in swarms of Culex pipiens pipiens (Diptera, Culicicdae). Parazitologia 33: 304–309. [In Russian]

Gad AM, Abdel Kader M, Farid HA, Hassan AN (1995) Absence of mating barriers between autogenous and anautogenous *Culex pipiens* L. in Egypt. Journal of Egyptian Society of Parasitology 25: 63–71.

Gordeev M, Goriacheva I, Shaikevich E, Ejov M (2004) Variability of the internal transcribed spacer of the ribosomal DNA among five Palearctic species of anopheline mosquitoes. European mosquito Bulletin 17: 14–19.

Harbach RE, Harrison BA, Gad AM (1984) *Culex* (*Culex*) *molestus* Forskal (Diptera, Culicidae): neotype designation, description, variation, and taxonomic status. Proceedings of the Entomological Society of Washington 86: 521–542. doi: 10.2987/8756-971X-28.4.10

Harbach RE, Dahl C, White GB (1985) *Culex* (*Culex*) *pipiens* Linnaeus (Diptera, Culicidae): concepts, type designations, and description. Proceedings of the Entomological Society of Washington 87: 1–24.

Harbach RE (2012) *Culex pipiens*: species versus species complex taxonomic history and perspective. Journal of the American Mosquito Control Association 28(4): 10–23.

Ivanov IO (1984) Swarming of *Culex pipiens* pipiens L. Meditsinskaia parazitologiia i parazitarnye bolezni 1: 22–25. [In Russian]

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Computer Applications in the Biosciences 8: 275–282. doi: 10.1093/bioinformatics/8.3.275

Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16: 111–120. doi: 10.1007/BF01731581

Konopka RJ, Benzer S (1971) Clock mutants of *Drosophila melanogaster*. Proceedings of the National Academy of Sciences of the United States of America 68(9): 2112–2116. doi: 10.1073/pnas.68.9.2112

Lopatin OE (2000) Allozyme polymorphism of the mosquitoes *C. p. pipiens*, *C. torrentium* and *C. vagans*. In: *Culex pipiens pipiens* mosquitoes: taxonomy, distribution, ecology, physiology, genetics, applied importance and control. Pensoft, Sofia-Moscow, 130–141.

Meuti ME, Stone M, Ikeno T, Denlinger DL (2015) Functional circadian clock genes are essential for the overwintering diapause of the Northern house mosquito, *Culex pipiens*. Journal of Experimental Biology 218: 412–422. doi: 10.1242/jeb.113233 218:412-422

Moriyama Y, Sakamoto T, Karpova SG, Matsumoto A, Noji S, Tomioka K (2008) RNA Interference of the Clock Gene period Disrupts Circadian Rhythms in the Cricket Gryllus bimaculatus. Journal of Biological Rhythms 23(4): 308–318. doi: 10.1177/0748730408320486

Nicolescu G, Linton Y-M, Vladimirescu A, Howard TM, Harbach RE (2004) Mosquitoes of the *Anopheles maculipennis* group (Diptera: Culicidae) in Romania, with the discovery and formal recognition of a new species based on molecular and morphological evidence. Bulletin of Entomological Research 94(6): 525–535. doi: 10.1079/BER2004330

Nudelman S, Galun R, Kitron U, Spielman A (1988) Physiological characteristics of *Culex pipiens* populations in the Middle East. Medical and Veterinary Entomology 2: 161–169. doi: 10.1111/j.1365-2915.1988.tb00066

Rivas GBS, Souza NA, Peixoto AA (2008) Analysis of the activity patterns of two sympatric sandfly siblings of the Lutzomyia longipalpis species complex from Brazil. Medical and Veterinary Entomology 22: 288–290. doi: 10.1111/j.1365-2915.2008.00742.x

Rona LD, Carvalho-Pinto CJ, Peixoto AA (2010) Molecular evidence for the occurrence of a new sibling species within the *Anopheles* (*Kerteszia*) *cruzii* complex in south-east Brazil. *Malaria* Journal 9: 33. doi: 10.1186/1475-2875-9-33

Sakai T, Ishida N (2001) Circadian rhythms of female mating activity governed by clock genes in *Drosophila*. Proceedings of the National Academy of Sciences of the United States of America 98: 9221–9225. doi: 10.1073/pnas.151443298

Sehgal A, Price J, Man B, Young M (1994) Loss of circadian behavioral rhythms and *per* RNA oscillations in the *Drosophila* mutant *timeless*. Science 263: 1603–1606. doi: 10.1126/science.8128246

Shaikevich E, Zakharov IA (2010) Polymorphism of mitochondrial *COI* and nuclear ribosomal ITS2 in *Culex pipiens* complex and in *Culex torrentium* (Diptera, Culicidae). Comparative Cytogenetics 4: 161–174. doi: 10.3897/compcytogen.v4i2.45

Shaikevich E, Zakharov IA (2014) Coevolution of Symbiotic Bacteria *Wolbachia* and Host mtDNA in Russian Populations of the *Culex pipiens* Mosquito Complex. Russian Journal of Genetics 50(11): 1234–1237. doi: 10.1134/S1022795414110131

Shinkawa Y, Takeda S, Tomioka K, Matsumoto A, Oda T, Chiba Y (1994) Variability in circadian activity patterns within the *Culex pipiens* complex (Diptera: Culicidae). J Med Entomol. 31(1): 49–56. doi: 10.1093/jmedent/31.1.49

Summa K, Urbanski JM, Zhao X, Poelchau M, Armbruster P (2012) Cloning and sequence analysis of the circadian clock genes period and timeless in *Aedes albopictus* (Diptera: Culicidae). Journal of Medical Entomology 49: 777–782. doi: 10.1603/ME11171

Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proceedings of the National Academy of Sciences USA 101: 11030–11035. doi: 10.1073/pnas.0404206101

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution 30: 2725–2729. doi: 10.1093/molbev/mst197

Tauber E, Roe H, Costa R, Hennessy JM, Kyriacou CP (2003) Temporal mating isolation driven by a behavioral gene in *Drosophila*. Current Biology 13: 140–145. doi: 10.1016/S0960-9822(03)00004-6

Tauber E, Zordan M, Sandrelli F, Pegoraro M, Osterwalder N, Breda C, Daga A, Selmin A, Monger K, Benna C, Rosato E, Kyriacou CP, Costa R (2007) Natural selection favors a newly derived timeless allele in *Drosophila melanogaster*. Science 316: 1895–8. doi: 10.1126/science.1138412

Vinogradova EB (2000) *Culex pipiens pipiens* mosquitoes: taxonomy, distribution, ecology, physiology, genetics, applied importance and control. Pensoft, Sofia-Moscow, 250 pp.

Vinogradova EB, Shaikevich EV (2007) Morphometric, physiological and molecular characteristics of underground populations of the urban mosquito *Culex pipiens* Linneus f. *molestus* Forskal (Diptera: Culicidae) from several areas of Russia. European Mosquito Bulletin 22: 17–24.

Weitzel T, Collado A, Jöst A, Pietsch K, Storch V, Becker N (2009) Genetic differentiation of populations within the *Culex pipiens* Complex (Diptera: Culicidae) and phylogeny of related species. Journal of the American Mosquito Control Association 25: 6–17. doi: 10.2987/08-5699.1

Supplementary material 1

Aligned nucleotide sequences of *per* gene.

Authors: Elena V. Shaikevich, Ludmila S. Karan, Marina V. Fyodorova

Data type: primary data

Explanation note: DNA sequences of three clones of each individual *C. pipiens* are presented and compared with sequences of *C. quinquefasciatus* (CPIJ007193) and *C. pipiens* from the USA (KM355980).

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 2

Aligned nucleotide sequences of *tim* gene.

Authors: Elena V. Shaikevich, Ludmila S. Karan, Marina V. Fyodorova

Data type: primary data

Explanation note: DNA sequences of three clones of each individual *C. pipiens* are presented and compared with sequences of *C. quinquefasciatus* (CPIJ007193) and *C. pipiens* from the USA (KM355980).

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 3

Analysis of the divergent between two forms of *C. pipiens* based on comparison of the exon 1 of the gene *tim* sequences.

Authors: Elena V. Shaikevich, Ludmila S. Karan, Marina V. Fyodorova

Data type: primary data

Explanation note: Nucleotide and amino acid sequences of the *tim* gene exon1 in compare with sequence of *C. quinquefasciatus*.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 4

Aligned *tim* nucleotide non-coding sequences.

Authors: Elena V. Shaikevich, Ludmila S. Karan, Marina V. Fyodorova

Data type: primary data

Explanation note: Intron's DNA sequences of individual *C. pipiens* f. "pipiens" and *C. pipiens* f. "molestus" are presented.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 5

Codon-based Test of Neutrality for analysis between *per* gene sequences of *C. pipiens* both forms.

Authors: Elena V. Shaikevich, Ludmila S. Karan, Marina V. Fyodorova

Data type: measurement of Z-test value

Explanation note: The test statistic (dN - dS) and the probability of rejecting the null hypothesis of strict-neutrality (dN = dS) are shown base on the differences between *per* gene sequences of *C. pipiens* both forms.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 6

Codon-based Test of Neutrality for analysis between *tim* gene sequences of *C. pipiens* both forms.

Authors: Elena V. Shaikevich, Ludmila S. Karan, Marina V. Fyodorova

Data type: measurement of Z-test value

Explanation note: The test statistic (dN - dS) and the probability of rejecting the null hypothesis of strict-neutrality (dN = dS) are shown base on the differences between *tim* gene sequences of *C. pipiens* both forms.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 7

Codon-based Test of of Purifying Selection for analysis between the exon1 of the gene *tim* sequences of *C. pipiens* both forms.

Authors: Elena V. Shaikevich, Ludmila S. Karan, Marina V. Fyodorova

Data type: measurement of Z-test value

Explanation note: The test statistic (dN - dS) and the probability of rejecting the null hypothesis of strict-neutrality (dN = dS) are shown base on the differences between the exon1 of the gene *tim* sequences of *C. pipiens* both forms.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.